

## WHAT IS CLAIMED IS:

1. A method of identifying a functional fragment pair in a protein, said method comprising:
  - 5 preparing fragments of a marker protein wherein each fragment has a break-point terminus within a solvent exposed loop of said marker protein, wherein the N or C terminal residue of each C or N terminal fragment, respectively, constitutes said break-point terminus, to obtain a marker fragment library;
  - expressing in a multiplicity of host cells, members of said marker fragment
  - 10 library;
  - isolating host cells expressing said marker protein as indicative of a cell containing a first member and a second member of a fragment pair which have formed a functionally reconstituted said marker protein, whereby said functional fragment pair is identified.
  - 15
2. The method according to Claim 1, wherein said functionally reconstituted marker protein confers a directly selectable signal.
3. The method according to Claim 1, wherein said first and said second member of  
20 said fragment pair together comprise one of a non-continuous, contiguous, or overlapping sequence of said marker protein and comprise between about 90 to 110% of the total length of said marker protein.
4. The method according to Claim 1, wherein said first member and said second  
25 member further each comprise a cysteine residue within 5 amino acid positions from said break-point terminus, so that a disulfide bond can form between said first member and said second member.
5. The method according to Claim 4, wherein said cysteine residue is at said  
30 break-point terminus.
6. The method according to Claim 1, wherein said protein is an enzyme.

7. The method according to Claim 5, wherein said enzyme is a  $\beta$ -lactamase.
8. The method according to Claim 1, wherein said fragments of said marker protein  
5 are each expressed as fusion proteins with one a fos or jun transcription factor.
9. A method of identifying a second oligopeptide to which a first oligopeptide binds,  
said method comprising:  
co-expressing in a multiplicity of host cells said first oligopeptide and said  
10 second oligopeptide wherein said second oligopeptide is encoded by a member of a  
library, each as a fusion protein with a first member and a second member of a  
fragment pair of a marker protein, respectively, obtained according to the method of  
Claim 1, wherein binding of said first oligopeptide to said second oligopeptide  
results in the functional reassembly of said marker protein;  
15 isolating host cells expressing said marker protein as indicative of a cell  
containing a first oligopeptide and a second oligopeptide which have interacted; and  
sequencing plasmids containing expression cassettes coding for said fusion  
proteins, whereby said second oligopeptide to which said first oligopeptide binds is  
identified.  
20
10. The method according to Claim 9, wherein each of said fusion proteins further  
comprises a signal peptide.
11. The method according to Claim 10, wherein said signal peptide provides for  
25 translocation to the periplasm of a bacterial cell.
12. The method according to Claim 11, wherein said first oligopeptide and said second  
oligopeptide are extracellular proteins.
- 30

13. The method according to Claim 10, wherein each of said fusion proteins further comprises a flexible polypeptide linker between said break-point terminus and said first or second oligonucleotide.
- 5 14. The method according to Claim 9, wherein said fusion protein further comprises at least one of the following:
- i) a randomly-encoded peptide of 3-12 amino acids between said break-point terminus and said flexible polypeptide linker;
  - ii) a cysteine residue within 5 amino acid positions from said break-point; and
  - 10 iii) 1-3 codon changes within said member of said fragment pair introduced by PCR amplification of a nucleotide sequence encoding for a member of said fragment pair under error-prone conditions, to enhance folding stability of a reconstituted marker protein.
- 15 15. The method according to Claim 9, further comprising a randomly-encoded peptide of 3-12 amino acids separately co-expressed as a fusion to the N-terminus of a thioredoxin.
16. The method according to Claim 9, wherein said host cell is an *E. coli* cell.
- 20 17. The method according to Claim 9, wherein said marker protein is an enzyme.
18. The method according to Claim 17, wherein said enzyme is a  $\beta$ -lactamase.
- 25 19. The method according to Claim 9, wherein said first oligopeptide is selected from the group consisting of a single chain antibody Fv fragment, an antibody light chain variable region, and a cell surface molecule, and said second oligopeptide is a randomly encoded peptide inserted into the active site of a thioredoxin or a phosphorylation-regulated signal transducer protein.
- 30 20. The method according to Claim 19, wherein said cell surface molecule is CD40.

21. The method according to Claim 19, wherein said phosphorylation-regulated signal transducer protein is a tyrosine kinase.

22. A fragment complementation system, said system comprising:

5 a first oligopeptide comprising an N-terminal fragment with a C-terminal break-point, and a second oligopeptide comprising a C-terminal fragment with a N-terminal break-point, wherein said N-terminal fragment and said C-terminal fragment each are derived from a marker protein and reassemble to form a functionally reconstituted marker protein.

10

23. The fragment complementation system according to Claim 22, wherein said first oligopeptide and said second oligopeptide each further comprise a cysteine residue within 5 amino acid positions of said break-point.

15

24. The method according to Claim 23, wherein said cysteine residue is at said break-point.

25. A fragment complementation system, said system comprising:

20

a first oligopeptide comprising an N-terminal fragment fused through a break-point to a flexible polypeptide linker and a first interactor domain; and  
a second oligopeptide comprising a second interactor domain and a flexible polypeptide linker fused through a break-point to a C-terminal fragment, wherein said N-terminal fragment and said C-terminal fragment are both derived from a marker protein with a directly selectable signal, and wherein said N-terminal fragment and said C-terminal fragment are obtained according to the method of

25

Claim 1, and wherein said N-terminal and said C-terminal fragment functionally reconstitute said marker protein only upon binding of said first interactor domain with said second interactor domain.

30

26. The fragment complementation system according to Claim 25, wherein said first and said second oligopeptide further comprise a signal peptide.

- 5
27. The fragment complementation system according to Claim 25, wherein said N-terminal and said C-terminal fragments together comprise one of a contiguous, overlapping or non-continuous sequence of said marker protein and comprise between about 90 to 110% of the total length of said marker protein.
28. The fragment complementation system according to Claim 27, wherein functional reconstitution of said marker protein is enhanced by introducing at least one of the following modifications to at least one of said first and said second oligopeptide sequences:
- 10 i) a randomly-encoded peptide of 3-12 amino acids encoded between said fragment and said flexible polypeptide linker,
- ii) a randomly-encoded peptide of 3-12 amino acids expressed separately and operably fused to the N-terminus of a thioredoxin,
- 15 iii) a cysteine residue encoded between said fragment and said flexible polypeptide linker, or
- iv) 1-3 codon changes per fragment molecule introduced by PCR amplifying a nucleotide sequence that encodes for said fragment under error-prone conditions to enable more stable folding of a reconstituted marker protein.
- 20 29. The fragment complementation system according to Claim 25, wherein said directly selectable signal is a visible phenotypic change or antibiotic resistance.
30. The fragment complementation system according to Claim 25, wherein said protein that has a directly selectable signal is an enzyme.
- 25
31. The fragment complementation system according to Claim 28, wherein said first interactor domain is selected from the group consisting of a single chain antibody Fv fragment, an antibody light chain variable region, and a cell surface molecule, and said second interactor domain comprises a randomly encoded peptide inserted
- 30 into the active site of *E. coli* thioredoxin or a phosphorylation-regulated signal transducer protein.

32. The fragment complementation system according to Claim 31, wherein said cell surface molecule is CD40.

33. The fragment complementation system according to Claim 31, wherein said phosphorylation-regulated signal transducer protein is a tyrosine kinase.

34. The fragment complementation system according to Claim 25, wherein said first interactor domain encodes a polypeptide from a first library and said second interactor domain encodes a polypeptide from a second library.

35. A fragment complementation system, said system comprising:  
a first oligopeptide comprising an N-terminal fragment of a  $\beta$ -lactamase fused through a break-point to a flexible polypeptide linker and a first interactor domain;  
and  
a second oligopeptide comprising a second interactor domain and a flexible polypeptide linker fused through a break-point to a C-terminal fragment of a  $\beta$ -lactamase, wherein said N-terminal and said C-terminal fragment functionally reconstitute said  $\beta$ -lactamase upon binding of said first interactor domain with said second interactor domain.

36. The fragment complementation system according to Claim 35, wherein functional reconstitution of said  $\beta$ -lactamase is enhanced by introducing at least one of the following modifications to at least one of said first and said second oligopeptide sequences:

- i) a randomly-encoded peptide of 3-12 amino acids encoded between said fragment and said flexible polypeptide linker,
- ii) a randomly-encoded peptide of 3-12 amino acids expressed separately and operably fused to the N-terminus of a thioredoxin,
- iii) a cysteine residue encoded between said fragment and said flexible polypeptide linker, or

iv) 1-3 codon changes per fragment molecule introduced by PCR-amplifying a nucleotide sequence that encodes for said fragment under error-prone conditions to enable more stable folding of a reconstituted marker protein.

5        37.    The fragment complementation system according to Claim 35, wherein said randomly-encoded peptide of 3-12 amino acids, is a tripeptide, and wherein a tripeptide fused to said N-terminal fragment is selected from the group consisting of HSE, NGR, GRE and EKR, and a tripeptide fused to said C-terminal fragment is selected from the group consisting of REQ, QGN, DGR GRR and GNS.

10

38.    The fragment complementation system according to Claim 36, wherein said break-point of said N-terminal fragment or said C-terminal fragment is within ten residues in either direction from a junction between amino acid residues selected from the group consisting of N52/S53, E63/E64, Q99/N100, P174/N175, E197/L198, K215/V216, A227/G228, and G253/K254.

15

39.    The fragment complementation system according to Claim 36, wherein said break-point of said N-terminal fragment or said C-terminal fragment is within ten residues in either direction of a junction between amino acid residues E197 and L198.

20

40.    The fragment complementation system according to Claim 39, wherein said randomly-encoded peptide of 3-12 amino acids, comprises the tripeptide GRE.

25

41.    The fragment complementation system according to Claim 35, wherein said N-terminal fragment comprises at least one mutation selected from the group consisting of K55E, P62S and M182T.

30

42.    An expression cassette comprising:  
as operably linked components in the direction of transcription nucleotide sequences encoding for:  
(i) a promoter functional in a host cell;  
(ii) a polypeptide interactor domain;

- (iii) a flexible polypeptide linker; and
- (iv) a C-terminal fragment of a marker protein that provides for a selectable phenotype.

- 5      43.      An expression cassette comprising:  
as operably linked components in the direction of transcription nucleotide sequences  
encoding for:
- (i) a promoter functional in a host cell;
  - (ii) an N-terminal fragment of a protein that provides for a selectable phenotype;
  - 10      (iii) a flexible polypeptide linker; and
  - (iv) a polypeptide interactor domain.

- 15      44.      The expression cassette according to Claim 42 or 43, further comprising a sequence  
encoding for a signal peptide.

- 20      45.      The expression cassette according to Claim 44, wherein said a signal peptide  
provides for translocation to the periplasm of a bacterial cell.

- 25      46.      The expression cassette according to Claim 45, wherein said interactor domain is an  
extracellular protein.

- 30      47.      The expression cassette according to Claim 42 or 43, wherein said marker protein  
that provides for a selectable phenotype is a  $\beta$ -lactamase.

- 35      48.      The expression cassette according to Claim 42, further comprising a sequence  
encoding for at least one of a randomly encoded peptide of from 3-12 amino acids  
or a cysteine residue operatively joined between said sequence encoding for said N-  
terminal fragment and said sequence encoding for said flexible polypeptide linker.

- 40      49.      The expression cassette according to Claim 43, further comprising a sequence  
encoding for at least one of a randomly encoded peptide of from 3-12 amino acids  
and a cysteine residue operatively joined between said sequence encoding for said



flexible polypeptide linker and said sequence encoding for said C-terminal fragment.

50. A host cell comprising a first and a second expression cassette, said first expression  
5 cassette according to Claim 42 and said second expression cassette according to Claim 43.

51. A method for identifying epitopes that bind to an immunoglobulin variable region,  
said method comprising:  
10 co-expressing from plasmids together in a host cell a first oligopeptide and a second oligopeptide, said first oligopeptide comprising an N-terminal fragment of  $\beta$ -lactamase fused operably in frame through a cysteine residue or a stabilizing tripeptide to a flexible polypeptide linker and a first interactor domain comprised of a randomly encoded peptide inserted into the active site of thioredoxin, and said  
15 second oligopeptide comprising a second interactor domain comprised of a single chain Fv fragment or an antibody light chain variable region and a flexible polypeptide linker fused operably in frame through a cysteine residue or a stabilizing tripeptide to a C-terminal fragment of  $\beta$ -lactamase, wherein the binding of said first interactor domain with said second interactor domain results in the  
20 functional reconstitution of said  $\beta$ -lactamase, and  
isolating host cells resistant to ampicillin; and  
sequencing plasmids containing expression cassettes coding for said first and second oligopeptides, whereby said epitopes that bind to said immunoglobulin variable regions are identified.

25

52. A method of identifying interactions between an extracellular domain of a transmembrane protein and a polypeptide, said method comprising:  
individually expressing from plasmids together in a host cell a first oligopeptide and a second oligopeptide, said first oligopeptide comprising an  
30 N-terminal fragment of  $\beta$ -lactamase fused operably in frame through a cysteine residue or a stabilizing tripeptide to a flexible polypeptide linker and a first interactor domain comprised of a randomly encoded peptide inserted into the active

site of thioredoxin, and said second oligopeptide comprising a second interactor domain comprised of a transmembrane protein and a flexible polypeptide linker fused operably in frame through a cysteine residue or a stabilizing tripeptide to a C-terminal fragment of  $\beta$ -lactamase, wherein the binding of said first interactor domain with said second interactor domain results in the functional reconstitution of said  $\beta$ -lactamase, and

isolating host cells resistant to ampicillin; and

sequencing plasmids containing expression cassettes coding for said first and second oligopeptides, whereby said polypeptide that binds to said transmembrane protein is identified.

53. The method according to Claim 52, wherein said transmembrane protein is an immune cell protein.

54. The method according to Claim 53, said immune cell protein is CD40.

55. A method for monitoring the occurrence of protein-protein interactions in a sample, said method comprising:

co-expressing in a host cell a first oligopeptide member of a first cellular library and a second oligopeptide member of a second cellular library, each as a fusion protein with a first member and a second member of a fragment pair of a marker protein, respectively, obtained according to the method of Claim 1, wherein binding of said first oligopeptide to said second oligopeptide results in the functional reassembly of said marker protein, and

isolating host cells expressing said marker protein as indicative of a cell containing a first member and a second member of a fragment pair which have functionally reconstituted said marker protein;

sequencing plasmids containing expression cassettes coding for said fusion proteins, whereby said protein-protein interactions are monitored.

56. A method for identifying oligopeptide interactions between two different proteomes, said method comprising:

co-expressing in a host cell a first oligopeptide member of a first cellular library and a second oligopeptide member of a second cellular library, each as a fusion protein with a first member and a second member of a fragment pair of  $\beta$ -lactamase, respectively, obtained according to the method of Claim 1, wherein  
5 binding of said first oligopeptide to said second oligopeptide results in the functional reassembly of said  $\beta$ -lactamase, and

isolating host cells resistant to ampicillin;

sequencing plasmids containing expression cassettes coding for said fusion proteins, whereby said oligopeptide interactions between two different proteomes  
10 are identified.

57. The method according to Claim 55 or 56, wherein said cellular library is from a tumor cell or an immune cell.

58. A method of high-throughput identification of compounds that inhibit  
15 phosphorylation-regulated cell signal transducers, said method comprising:

co-expressing from plasmids together in a host cell a first oligopeptide and a second oligopeptide, said first oligopeptide comprising an N-terminal fragment of  $\beta$ -lactamase fused operably in frame through a cysteine residue or a stabilizing  
20 tripeptide to a flexible polypeptide linker and a first interactor domain comprised of a single chain Fv fragment or an antibody light chain variable region that binds a nonphosphorylated active site of a phosphorylation-regulated cell signal transducer, and said second oligopeptide comprising a second interactor domain comprised of a phosphorylation-regulated cell signal transducer protein and a flexible polypeptide  
25 linker fused operably in frame through a cysteine residue or a stabilizing tripeptide to a C-terminal fragment of  $\beta$ -lactamase, wherein the binding of said first interactor domain with said second interactor domain results in the functional reconstitution of said  $\beta$ -lactamase, and

identifying said compounds that result in a host cell turning color in the  
30 presence of chromogenic  $\beta$ -lactamase substrate.

59. The method according to Claim 58, wherein said phosphorylation-regulated cell signal transducer protein is a tyrosine kinase.

60. The method according to Claim 59, wherein said tyrosine kinase is Her-2/neu.

5

61. An enzyme complementation system to select for simultaneous incorporation of multiple genetic elements into a host cell, said system comprising:

co-expressing in a host cell an N-terminal fragment and a C-terminal fragment of an antibiotic resistance protein, wherein said N-terminal fragment expresses from a first recombinant sequence also encoding for a first trait, and said C-terminal fragment expresses from a second recombinant sequence also encoding for a second trait, wherein said cell expressing polypeptide from both said first and said second recombinant sequence produces said N-terminal fragment and said C-terminal fragment in a sufficient amount to reconstitute said antibiotic resistance protein, and

15

ii) isolating cells resistant to said antibiotic.

62. A method of activating a  $\beta$ -lactam derivative of an anti-tumor compound in a host in need thereof, said method comprising:

20

i) simultaneously administering to said host a first oligopeptide and a second oligopeptide, said first oligopeptide comprising an N-terminal fragment of  $\beta$ -lactamase, a flexible polypeptide linker and a first single chain Fv fragment against an epitope of a tumor protein, said second oligopeptide comprising a second single chain Fv against a second non-overlapping epitope of said tumor protein, a flexible polypeptide linker and a C-terminal fragment of  $\beta$ -lactamase, wherein said single chain Fv fragments bind to said epitopes resulting in the functional reconstitution of  $\beta$ -lactamase, and

25

ii) administering said  $\beta$ -lactam derivative of said anti-tumor compound to said host, whereby said derivative is activated by said reconstituted  $\beta$ -lactamase near said tumor protein.

30